Beilstein-Institut ESCEC, Oct. 5th - 8th 2003, Rüdesheim, Germany

METHODS OF DESIGN OF OPTIMAL EXPERIMENTS WITH APPLICATION TO PARAMETER ESTIMATION IN ENZYME CATALYTIC PROCESSES

HANS GEORG BOCK, STEFAN KÖRKEL, EKATERINA KOSTINA* AND JOHANNES P. SCHLÖDER

Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg,
Im Neuenheimer Feld 368, 69120 Heidelberg, Germany
E-Mail: *ekaterina.kostina@iwr.uni-heidelberg.de

Received: 20th April 2004 / Published: 1st October 2004

ABSTRACT

This paper deals with the identification of kinetic parameters in enzyme catalytic processes. Experience shows that the experiments performed do no deliver measurement data sufficient for the identification of parameters. New optimal experiments are needed. We suggest effective algorithms and software for parameter estimation and design of optimal experiments, based on multiple shooting and specially tailored, structure-exploiting reduced Gauss-Newton and SQP methods. The methods are applied to optimal experimental design and parameter estimation in enzyme catalytic processes.

INTRODUCTION

We consider the problem of identifying enzyme operating stability which is very important e.g. in industrial production of enzymes or in understanding enzymatic pathways in a living system. Traditionally enzyme stability is determined empirically by time-consuming experiments: under constant reaction conditions (temperatures) the half-life of the catalyst is measured [1]. This procedure is repeated under different temperatures. Another approach for faster identification of enzyme operating stability is suggested in [2]. In an instationary operated reactor the reaction temperature is increased linearly and parameters of a kinetic model based on Arrhenius equations are estimated. Then the stability properties of enzymes are estimated as functions of the kinetic parameters. Experience shows however that such experiments do not deliver enough data to identify the kinetic parameters reliably.

Optimal experiments (temperature profiles) are necessary that allow the estimation of the kinetic parameters of enzyme catalytic processes and thus stability properties. This paper focuses on numerical methods for parameter estimation and design of optimal experiments.

The paper is organized as follows. Firstly, we describe parameter estimation problems in ordinary differential equations as well as specifying the model of enzyme reaction under consideration. Secondly, the method used for parameter estimation is described. It follows a very successful solution approach, namely the boundary value problem (BVP) approach, which combines infeasible point optimization algorithms like the generalized Gauss-Newton method with BVP techniques like multiple shooting to discretize the dynamic system. The discretized equations are then treated as nonlinear equality constraints. We solve the resulting finite-dimensional, possibly large-scale, nonlinear constrained least squares problem with the generalized Gauss-Newton method. Under certain regularity assumptions the method shows the final linear rate of local convergence. To ensure global convergence, a newly developed efficient step-size strategy is used. Special attention is paid to the discussion of why the generalized Gauss-Newton method is particularly appropriate for parameter estimation and does not converge to solutions with large residuals. The quality of parameter estimates is defined by statistical analysis, and the results of parameter estimation for an immobilized enzyme system finish this part of the paper. Thirdly, theoretical justification methods for the design of optimal experiments as well as numerical and experimental results are presented. Optimal experiments are designed for the nominal values of parameters which are known only to lie in a confidence region. Fourthly, the design of experiments that are less sensitive to parameter uncertainty is described. The paper is finished by conclusions.

PARAMETER ESTIMATION PROBLEM FOR ENZYME CATALYTIC PROCESSES

Parameter estimation problems in dynamic processes

The problem of identification of unknown parameters in dynamic models is among the most important tasks in mathematical modelling of dynamic processes. It can be described as follows. Let the dynamics of the model be described by a system of ordinary differential equations

$$\dot{y} = f(t, y(t), p, q, u(t)), t \in [t_0, t_{end}]$$

where the right-hand side *f* depends on an vector of unknown parameters $p \in R^{n_p}$, given control functions $u: [t_0;t_{end}] \to R^{n_u}$ and given control variables $q \in R^{n_q}$. It is assumed, that at the given times $t_j, j = 1,..., M,$ measurements $\eta_{ij}, i = 1,..., M_j, j = 1,..., M$, of the observation functions bij are available

$$\eta_{ij} = b_{ij}(y(t_j), p^{true}, q) + \varepsilon_{ij}$$

which are subject to measurement errors ε_{ij}. Here p^{true} are the "true" values of the parameters. Note, that several model quantities can be measured at a time t_j. According to the common approach, in order to determine the unknown parameters an optimization problem is solved. The possible constraints of this problem describe the specifics of the model (constraints on the initial and terminal states, constraints on parameters, etc.) and can be formally written as constraints at times $\theta_1, \ldots, \theta_\kappa \in [t_0, t_{end}]$

$$r_{con}(y(\theta_1), \ldots y(\theta_\kappa), p, q) = 0$$

As an objective functional in the optimization problem, typically a norm of the measurement errors is used. The type of the norm is motivated by the statistical distribution of the measurement errors. If the errors are independent, normally distributed with zero mean and known variances $(N(0, \sigma_{ij}^2))$, minimizing a weighted least squares function

$$\min \sum_{i,j} (\eta_{ij} - b_{ij}(y(t_j), p, q))^2 / \sigma_{ij}^2$$

yields a maximum likelihood estimate.

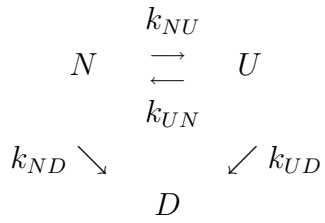
Summing up, mathematically the problems of parameter estimation can be written as follows:

Minimize the deviation of model response to measurement values such that the dynamics and the inital conditions of the dynamic process are fulfilled and possible further constraints are satisfied:

$$\begin{aligned} \min_{y,p} \quad & \sum_{j=1}^{\mathcal{M}} \sum_{i=1}^{\mathcal{M}_j} (\eta_{ij} - b_{ij}(y(t_j), p, q))^2 / \sigma_{ij}^2 \quad (1) \\ \text{s.t.} \quad & \dot{y}(t) = f(t, y(t), p, q, u(t)), t \in [t_0, t_{end}] \\ & r_{con}(y(\theta_1), ..., y(\theta_K), p, q) = 0. \end{aligned} \tag{1}$$

Model for enzyme catalytic processes

We consider parameter estimation for the catalytic reaction of enzymes [3]. The native (N), unfolded (U), and deactivated (D) enzymes are involved in the reaction which takes place in a reactor with a continuous inflow and a corresponding outflow. The reaction scheme is the following



where k_{NU}, k_{UN}, k_{UD}, k_{ND} denote rates of the corresponding reactions. Additionally a substrate S is involved, which is contained in the inflow and is degraded depending on the concentration of native enzymes.

Our mathematical model of this reaction scheme is based on the following assumptions:

- The concentrations C_N, C_U, C_D of the native, the unfolded and the deactivated enzymes respectively satisfy a conservation law:

$$C_N + C_U + C_D = C_{E_0}$$

where C_{E0} denotes an initial amount of active enzymes.

- Native and unfolded enzymes are in a quasi-steady-state:

$$\frac{C_U}{C_N} = K_U$$

with some constant K_U.

- A biochemical reaction involving the substrate S takes place. This reaction is described by the Michaelis-Menten kinetics:

$$r = r_{max}\frac{C_S}{k_m + C_S}, r_{max} = k_r C_N$$

where C_S is the concentration of the substrate, k_m is the Michaelis-Menten constant, r_{max} is the maximal rate of the biochemical reaction, k_r is a proportionality coefficient depending on the temperature T.

- The reactions rates k_{NU}, k_{UN}, k_{UD}, k_{ND}, as well as the steady state constant K_U and the proportionality factor k_r depend on the temperature T according to the Arrhenius law:

$$k = k_1 \exp\left(\frac{-k_2}{RT}\right) \qquad (2)$$

where R is the gas constant.

In particular:

$$k_{ND} = k_N^0 \exp\left(\frac{-\Delta h_N^*}{RT}\right),\ k_{UD} = k_d^0 \exp\left(\frac{-\Delta h_u^*}{RT}\right),$$
$$K_U = \exp\left(\frac{-\Delta h_u^0}{RT}\right) \exp\left(\frac{\Delta S_u^0}{R}\right),\ k_r = A \exp\left(\frac{-\Delta h_E^*}{RT}\right),$$

where

- A denotes the activation constant,
- Δh_E^* denotes the activation enthalpy,
- ΔS_u^0 denotes the deactivation entropy,
- Δh_u^0 denotes the deactivation enthalpy,
- k_d^0 denotes the decay constant,
- Δh_u^* denotes the activation enthalpy,
- k_N^0 denotes the native enzyme decay constant,
- Δh_N^* denotes the native enzyme activation enthalpy.

- The change of the substrate concentration C_S due to the inflow and the outflow of the substrate is defined by

$$\frac{\dot{V}}{V}\left(C_S^0 - C_S\right),$$

where V ist the volume of the reactor, $\dot{V}$ is the rate of the substrate inflow, C_S^0 is the inflow concentration of the substrate.

Taking into account these assumptions one can derive the mathematical model for the enzyme catalytic reaction

$$\begin{aligned}
\frac{d\,C_D}{d\,t} &= \left(k_d^0 \exp\left(\frac{-\Delta h_u^*}{RT}\right) K_U + \right. \\
&\quad \left. k_N^0 \exp\left(\frac{-\Delta h_N^*}{RT}\right)\right)(C_{E_0} - C_D)\,\frac{1}{1+K_U}, \quad C_D(0) = 0, \qquad (3)\\
\frac{d\,C_S}{d\,t} &= \frac{\dot{V}}{V}\left(C_S^0 - C_S\right) - r_{max}\frac{C_S}{k_m + C_S}, \quad C_S(0) = C_S^0.
\end{aligned}$$

with

$$K_U = \exp\left(\frac{-\Delta h_u^0}{RT}\right)\exp\left(\frac{\Delta S_u^0}{R}\right),\ r_{max} = A\exp\left(\frac{-\Delta h_E^*}{RT}\right)\frac{C_{E_0} - C_D}{1+K_U}.$$

Summing up, the catalytic reaction is modelled by a system of two differential equations (3) where

- the state variables are the concentration of the deactivated enzyme C_D and the concentration of the substrate C_S,
- the parameters to be estimated are

$$\begin{array}{llll} p_1 = \ln A, & p_2 = \Delta h_E^*, & p_3 = \Delta S_u^0, & p_4 = \Delta h_u^0, \\ p_5 = k_d^0, & p_6 = \Delta h_u^*, & p_7 = k_N^0, & p_8 = \Delta h_N^*. \end{array}$$

The process may be controlled by the temperature, this means that the control function u is a temperature profile T.

For numerical stability and better scaling we reformulate the Arrhenius kinetic terms (2) as follows:

$$k_1 \exp\left(-\frac{k_2}{RT}\right) \;=\; \exp\left(k_{1new}\frac{T^{-1}-T_0^{-1}}{T_1^{-1}-T_0^{-1}} + k_{2new}\frac{T_1^{-1}-T^{-1}}{T_1^{-1}-T_0^{-1}}\right),$$

where T_1 and T_0 are some temperature values to be chosen by the use, e.g. maximal and minimal temperature used in the experiments. Physically the coefficients k_{1new} and k_{2new} describe the rate of the corresponding reaction at T_1 and T_0 respectively.

We use this transformation for parameters describing reactions $N \to D$ and $U \to D$:

- instead of the parameters p_5 and p_6 we introduce the paramters p_{5new} and p_{6new} according to the formulae:

$$k_d^0 \exp\left(-\frac{\Delta h_u^*}{RT_0}\right) \;=\; \exp(p_{6new}),\; k_d^0 \exp\left(-\frac{\Delta h_u^*}{RT_1}\right) = \exp(p_{5new}),$$

- and analogously instead of the parameters p_7 and p_8, we introduce the parameters p_{7new} and p_{8new} according to the formulae:

$$k_N^0 \exp\left(-\frac{\Delta h_N^*}{RT_0}\right) \;=\; \exp(p_{8new}),\; k_N^0 \exp\left(-\frac{\Delta h_N^*}{RT_1}\right) = \exp(p_{7new}).$$

For the reader's convenience, we give the explicit relation between "old" and "new" parameters:

’new’ parameters

$$p_{5new} = \ln(k_d^0) - \tfrac{\Delta h_u^*}{RT_1},\quad p_{6new} = \ln(k_d^0) - \tfrac{\Delta h_u^*}{RT_0},$$

$$p_{7new} = \ln(k_N^0) - \tfrac{\Delta h_N^*}{RT_1},\; p_{8new} = \ln(k_N^0) - \tfrac{\Delta h_N^*}{RT_0};$$

'old' parameters

$$k_d^0 = \exp\left(\frac{T_1 p_{5new} - T_0 p_{6new}}{T_1 - T_0}\right), \quad \Delta h_u^* = \frac{RT_1T_0(p_{5new} - p_{6new})}{T_1 - T_0},$$

$$k_N^0 = \exp\left(\frac{T_1 p_{7new} - T_0 p_{8new}}{T_1 - T_0}\right), \quad \Delta h_N^* = \frac{RT_1T_0(p_{7new} - p_{8new})}{T_1 - T_0}.$$

Half-life

The important stability feature of enzymes is characterized by half-life [2].

Half-life (HL) is a time required to reduce the amount of a native enzyme to a half of the initial amount at a constant temperature.

Mathematically half-life is given by

$$\mathrm{HL}(T) \;=\; \ln 2(1 + K_U) / \left(K_U k_d^0 \exp\left(\frac{-\Delta h_u^*}{RT}\right) + k_N^0 \exp\left(\frac{-\Delta h_N^*}{RT}\right)\right).$$

Measurement function

The observation function - velocity of consumption of base necessary to neutralize the acidic reaction product - is given by the dosage of base

$$b := \frac{\dot{V}(C_S^0 - C_S(t))}{B_0}$$

where B_0 is the given concentration of base. Note that in the model under consideration we measure one quantity at a time.

BOUNDARY VALUE PROBLEM METHODS FOR PARAMETER ESTIMATION

A typical solution approach to parameter estimation which is found very often in practice is the initial value or single shooting approach: the ODE system is repeatedly solved as an initial value problem, and unknown parameters including possibly initial values are iteratively improved by some optimization procedure.

In contrast to that, our numerical solution of the parameter estimation problem is based on the Boundary Value Problem (BVP) approach going back to [4]. The basic idea consists in parameterizing the dynamic equations (initial or boundary value problem) like a boundary value problem (e.g., by multiple shooting) and then performing simultaneously (in one iteration loop) the minimization of the cost function and the fulfilment of the constraints given by the discretized boundary value problem. It has been shown [4,5], that BVP methods (based on multiple shooting or collocation) are much more stable and efficient than the single shooting approach when solving parameter estimation problems.

Multiple shooting

The scheme of multiple shooting consists of the following. First one chooses a suitable grid of multiple shooting nodes τ_j

$$t_0 = \tau_0 < \tau_1 < \ldots < \tau_m = t_{end}$$

covering the interval where measurements are given.

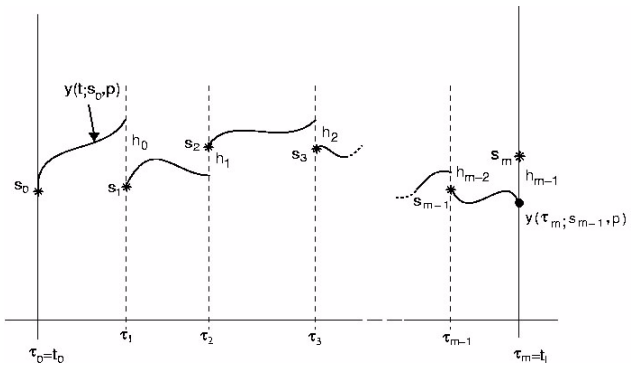


Figure 1. Multiple shooting approach.

At each grid point the values of the state variables s_j are chosen as additional unknowns and m ODE initial value problems

$$\dot{y} = f(t, y, p, q, u), \quad y(\tau_j) = s_j$$

are solved on each subinterval $I_j := [\tau_j, \tau_j + 1]$ to yield a solution $y(t; sj, p, q, u)$ for $t \in I_j$. Solutions of dynamic systems, generated by this procedure, are usually not continuous at τ_j. This has to be enforced by additional matching conditions. Inserting the computed values $y(t, s_j, p, q, u), \quad \tau_j \le t \le \tau_{1+1}$ into problem (1) one obtains a constrained optimization problem in the variables $(s, p) := (s_0,...,s_m, p)$

$$\begin{aligned} &\sum_{j=1}^{\mathcal{M}} \sum_{i=1}^{\mathcal{M}_j} (\eta_{ij} - b_{ij}(y(t_j), p, q))^2 / \sigma_{ij}^2, \\ &h_j(s_j, s_{j+1}, p) := y(\tau_{j+1}; s_j, p, q, u) - s_{j+1} = 0, \ j = 1, .., m-1, \\ &r_{con}(y(\theta_1), ..., y(\theta_{\mathcal{K}}), p, q) = 0. \end{aligned} \tag{4}$$

Multiple shooting possesses several advantages:

1. It is possible to include a priori information about the state variables, e.g. from the measurements or from expert knowledge, by a proper choice of initial guesses for the additional variables s_j. Thus, it can be ensured that the initial solutions $y(t;s_j, p, u)$ remain close to the observed data. It can be shown that this damps the influence of poor parameter guesses.

2. The adequate choice of initial guesses for the state variables (and the application of a Gauss-Newton method for the solution of the constrained least squares problem) typically avoids convergence to local minima with large residuals.

3. The scheme is numerically stable. The splitting of the integration interval limits error propagation and makes it possible to solve parameter estimation problems even for unstable or chaotic systems.

4. The BVP discretization induces a very specific structure in the problem equations, which can be exploited in particular for parallelization.

Gauss-Newton method

Parametrization of the dynamics yields a finite dimensional, possibly large-scale, nonlinear constrained least squares problem which can be formally written as

$$\begin{aligned} \min_{x \in R^n} \quad & ||F_1(x)||_2^2, \\ \text{s.t.} \quad & F_2(x) = 0. \end{aligned}$$

Note, that the equalities $F_2(\chi) = 0$ include the matching conditions induced by multiple shooting. We assume that the functions $F_i : D \subset R^n \to R^{m_i}$, $i = 1, 2$, are twice continuously differentiable. The number of variables in problem (5) is equal to number of differential equations multiplied by the number of multiple shooting nodes plus the number of parameters. To solve problem (5) we use a generalized Gauss-Newton method according to which a new iterate is (basically) generated by

$$x^{k+1} = x^k + t^k \Delta x^k, 0 < t_k \leq 1, \tag{6}$$

where the increment $\Delta\chi$ is the solution of the following linear constraint l_2 problem at $\chi = \chi^k$

$$\begin{aligned} \min_{\Delta x \in R^n} \quad & ||F_1(x) + J_1(x)\Delta x||_2^2, \\ \text{s.t.} \quad & F_2(x) + J_2(x)\Delta x = 0. \end{aligned} \tag{7}$$

Here, $J_1(\chi)$ and $J_2(\chi)$ denote the Jacobians of $F_1(\chi)$ and $F_2(\chi)$ respectively.

$$\text{rank} J_2(x) = m_2, \ \text{rank } J = n, \ J = J(x) = \begin{pmatrix} J_1(x) \\ J_2(x) \end{pmatrix}$$

then a linearized problem (7) has a unique solution $\Delta\chi^k$ and a unique Lagrange vector λ^k satisfying the following optimality conditions

$$\begin{array}{rll} J_1^T(x)J_1(x)\Delta x^k & -J_2^T(x)\lambda^k & = -J_1^T(x)F_1(x), \\ J_2(x)\Delta x^k & & = -F_2(x). \end{array} \qquad (8)$$

Using (8) one can easily show that $\Delta\chi^k$ can be formally written with the help of a solution operator J^+

$$\Delta x^k = -J^+(x^k)F(x^k),\ F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \end{pmatrix}.$$

The solution operator J^+ is a generalized inverse, that is it satisfies $J^+JJ^+ = J$, and is explicitly given by

$$J^+(x) = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} J_1^T(x)J_1(x) & J_2(x)^T \\ J_2(x) & 0 \end{pmatrix}^{-1} \begin{pmatrix} J_1(x)^T & 0 \\ 0 & I \end{pmatrix}.$$

Choosing the step lenght t^k by means of classical line search methods based on the exact penalty function

$$T_1(x) := ||F_1(x)||_2^2 + \sum_{i=1}^{m_2} \alpha_i |F_{2i}(x)|$$

with sufficiently large weights $\alpha_i < 0$, $i = 1,..., m_2$, ensures global convergence. However, it is well known that already in mildly ill-conditioned problems such a step-size strategy may be very inefficient since it may produce very small step-sizes. Therefore we use the "restrictive monotonicity test", see [5,6], that has proved to be very effective in practical applications.

Note, that the generalized Gauss-Newton method (6) - (7) has several advantages. First it does not use second order derivative information and the local linearized problems are linear constrained least squares problems. Under certain regularity assumptions at the solution, the method shows a good linear rate of local convergence. There are problems, however, for which the Gauss-Newton method may have a rather slow local convergence rate or may even fail. The reason is that the linearized model (7), which forms the basis of the Gauss-Newton method, is an inadequate representation of such nonlinear problems, since the second-order information cannot be ignored. These problems are called problems with large residuals.

Using SQP-type methods for the nonlinear constrained l_2 parameter estimation problem one could force convergence to a solution even in such a case. However, such solutions are undesirable in a certain sense. We can show that a solution of this type, even if it is a strict minimum of the nonlinear constrained l_2 problem (5), cannot be expected to be a continuous deformation of the "true" parameter values under perturbations caused by the measurement errors. Thus, slow local convergence of the full-step ($t^k \equiv 1$) Gauss-Newton indicates deficiencies in the model or lack of data and can be considered as an advantage of the method. For a detailed analysis see [5].

Solving the linear l_2 problem

At each iteration of a Gauss-Newton method a linear least squares problem (7) has to be solved. First, we make use of the special block structure of the Jacobian $J(\chi)$ which is induced by multiple shooting:

$$J = \left[\begin{array}{cccccc|c} D_0 & D_1 & & \dots & & D_m & D^p \\ \hline G_0^l & G_0^r & & & & & G_0^p \\ & \ddots & \ddots & & 0 & & \vdots \\ & & \ddots & \ddots & & & \vdots \\ & 0 & & \ddots & \ddots & & \vdots \\ & & & & G_{m-1}^l & G_{m-1}^r & G_{m-1}^p \end{array}\right].$$

Every block column corresponds to the derivatives with respect to the discretization variables and parameters in one subinterval. The block rows with G-matrices are the derivatives of the continuity conditions

$$G_j^l := \partial h_j / \partial s_j \quad G_j^r := \partial h_j / \partial s_{j+1} = -I, G_j^p := \partial h_j / \partial p.$$

The block rows with D-matrices correspond to the derivatives of the functions F_i of the cost functional and the constraints of the nonlinear problem (5) excluding the continuity conditions. We use a fast, stable and efficient structure exploiting decomposition (see [4, 5]) to reduce the large linear least squares problem to a linear least squares problem with smaller dimension.

The number of variables in the resulting problem is equal to the number of parameters plus the number of differential equations. This so-called condensed problem may be solved using the methods described in [7].

Statistical sensitivity analysis for the estimates

An important question in parameter estimation is how good the computed estimates are. The answer is provided by sensitivity analysis. If the experimental data is normally distributed then the estimated solution χ^* of the parameter estimation problem is also a random variable which is normally distributed in the first order $\chi^* \sim N(\chi^{true}, C)$ with the (unknown) true value χ^{true} as expected value and the variance-covariance matrix C given by

$$C = C(x,q,u) = J^{+}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} J^{+T}. \qquad (9)$$

The variance-covariance matrix describes the confidence ellipsoid which is an approximation of the nonlinear confidence region of the estimated variables. The 100(1- a)% linearized confidence ellipsoid $(0 \le \alpha \le 1)$ can be described by (see [5])

$$G_L(\alpha, x^{true}, q, u) = \left\{ x : x = x^{true} + J^{+}\begin{pmatrix} \eta \\ 0 \end{pmatrix}, ||\eta||_2^2 \le \gamma^2(\alpha) \right\}.$$

Here, the probability factor $\gamma^2(\alpha) = \chi^2_{n \angle m_2}\ (1 \angle \alpha)$ where $\chi^2_{n \angle m_2}\ (1 \angle \alpha)$ is the quantile of the χ^2 distribution. The linearized confidence ellipsoid $G_L(\alpha,\chi^{true},q,u)$ is contained exactly in a box determined by the confidence intervals

$$G_L(\alpha, x^{true}, q, u) \subset \bigtimes_{i=1}^{n} [x_i^{true} - \theta_i, x_i^{true} + \theta_i],$$

where $\theta_i = \sqrt{C_{ii}}\gamma(\alpha)$. Here, C_{ii} denotes the diagonal elements of the covariance matrix C. The values $\sqrt{C_{ii}}$ are known as standard deviations of the variables χ_i.

Results of parameter estimation for *Candida antarctica* lipase on ionic resin ("Novozym")

Applying the methods of parameter estimation described earlier, "Novozym" shows that the data derived from the initial experiment does not deliver enough information to identify all parameters, see Table 1. The settings of the initial experiment are defined by the temperature shown in Fig. 2. The parameter estimation results show that the information received in the experiment is not at all enough to estimate parameters reliably. Thus we need to design additional experiments in order to provide sufficiently good data. The theoretical justification and methods for design of optimal experiments are briefly described in the next section.

Table 1. Estimated values of parameters ± standard deviation after parameter estimation.

	Initial temperature profile
p_1	27.86 ± 4.42
p_2	48.98 ± 10.92
p_3	$1.73 \pm 2.39 \times 10^5$
p_4	$634.20 \pm 806.00 \times 10^6$
p_5	$-1.43 \pm 1.50 \times 10^7$
p_6	$-7.50 \pm 4.16 \times 10^7$
p_7	-4.15 ± 0.091
p_8	-8.63 ± 2.00

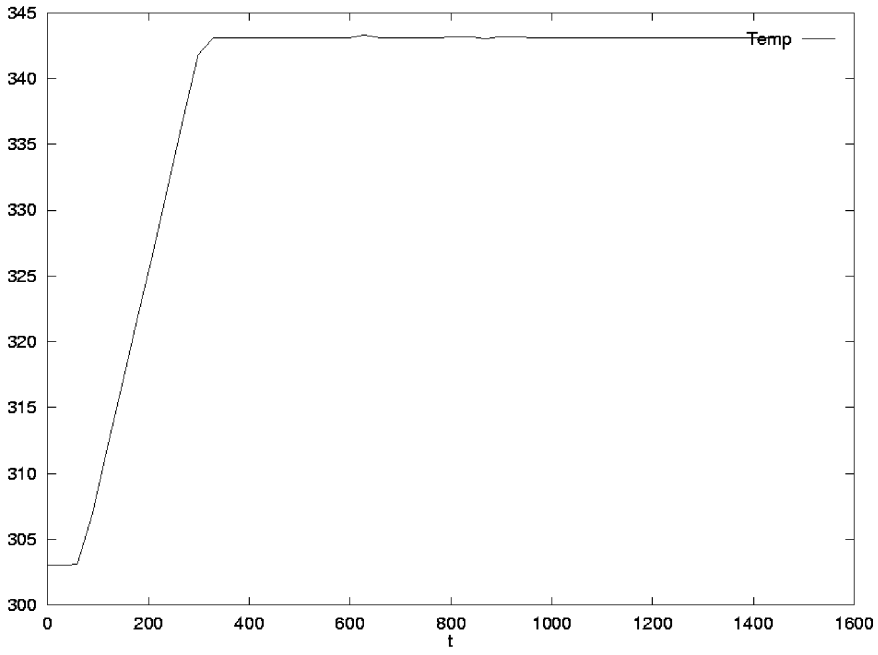


Figure 2. Initial temperature profile

OPTIMUM EXPERIMENTAL DESIGN: PROBLEM STATEMENT AND NUMERICAL METHODS

In this section we consider the problem of nonlinear optimum experimental design and discuss a solution approach to this problem.

Experimental design optimization problem

Since the experimental data is randomly distributed, the estimated parameters are also random variables. Data evaluated under different experimental conditions leads to estimations of parameters with good or poor confidence regions depending on the experimental conditions. We would like to find those experiments that result in the best statistical quality for the estimated parameters and at the same time satisfy additional constraints e.g. experimental costs, safety, feasibility of experiments, validity of the model etc.

The aim of optimum experimental design is to construct N_{ex} experiments by choosing appropriate experimental variables
$q = (q_1,...,q_{Nex})$, and experimental controls $u = (u_1,..., u_{Nex})$ in order to maximize the statistical reliability of the unknown variables under estimation. For this purpose we optimize a design criterion depending on the variance-covariance matrix C (9).

The optimum experimental design approach leads to an optimal control problem in ODE systems of the following form

$$\min_{q,u} \quad \Phi(C(y,q,u)) \tag{10}$$

$$\text{s.t.} \quad c_i(t, y_i(t), p, q_i, u_i(t)) \geq 0,\ i = 1,2,...,N_{ex}, \tag{11}$$

$$\dot{y}_i = f_i(t, y_i, p, q_i, u_i),\ i = 1,2,...,N_{ex},$$

$$r_{con,i}(y_i(\theta_1^i),...,y_i(\theta_{K\rangle}^i), p, q_i) = 0,\ i = 1,2,...,N_{ex}. \tag{12}$$

The constraint (11) describes control and path constraints for each experiment of N_{ex} experiments. Free variables of the optimization problem are the control profiles $u_i(t)$ (e.g. temperature profiles of cooling/heating) and the time-independent control variables q_i (e.g. initial concentrations, properties of the experimental device) for all experiments.

Since the "size" of a confidence region is described by variance-covariance matrix C any suitable function of matrix C has to be taken as a cost functional. Typical choices are

$$\begin{aligned} \Phi(C) &= \text{trace}(C), \\ \Phi(C) &= \lambda_{max}(C), \quad \text{where } \lambda_{max} \text{ denotes the largest eigenvalue of } C \\ \Phi(C) &= \max_i c_{ii}. \end{aligned}$$

Numerical methods

The experimental design optimization problem is a nonlinear constrained optimal control problem. The main difficulty lies in the non-standard objective function which is nonseparable and implicitly defined on the sensitivities of the underlying parameter estimation problem, i.e. on the derivatives of the solution of the ODE system with respect to the parameters and initial values.

The numerical methods are based on the direct approach, according to which the control functions are parametrized on an appropriate grid by support functions, the solution of the ODE systems and the state constraints are discretized. As a result we obtain a finite-dimensional constrained nonlinear optimization problem which is solved by an SQP method. The main effort for the solution of the optimization problem by the SQP method is spent on the calculation of the values of the objective function and the constraints as well as its gradients.

Efficient methods for derivative computations combining internal numerical differentiation [5] of the integration method and automatic differentiation of the model functions [8] have been developed, see [9,10,11]. For more detailed discussion of the numerical methods for nonlinear optimum experimental design see [11,12].

Results of parameter estimation for *Candida antarctica* lipase on ionic resin ("Novozym") with data from additional experiments

Using the methods of optimum experimental design we have optimized 5 additional temperature profiles. The following set-up has been chosen for the experiments: the duration of each experiment is 20 h; measurements are taken every half-hour; the temperature profile is parametrized as follows

$$T_i(t) = T_i + \text{slope}_i(t - t_i), t \in [t_i, t_{i+1}], T_i(t_{i+1}) = T_{i+1}, t_i = 60i, i = 0, ..., 19, t_{20} = 1200.$$

Additionally, there are the following restrictions on the design parameters T_i and slope_i.

$$293 \leq T_i \leq 343, \quad i = 0, ..., 20, \ -0.12 \leq \text{slope}_i \leq 0.25, \quad i = 0, ..., 19,$$

Results of parameter estimation for "Novozym" with data from the initial and additional experiments are presented in Tables 2 and 3. Table 3 shows the estimated values of half-life over an interesting temperature range. The optimal temperature profiles are presented in Fig. 3. Figure 4 presents the fits before and after parameter estimation. The results of parameter estimation show that now we can estimate all half-lives with standard deviation below 30 %.

Table 2. Estimated values of parameters ± standard deviation after parameter estimation.

	Initial profile	**5 designed + 1 initial profile**
p_1	27.86 ± 4.42	27.69 ± 0.81
p_2	48.98 ± 10.92	49.96 ± 1.97
p_3	1.73 ± 2.39 x 10^5	0.548 ± 0.075
p_4	634.20 ± 806.00 x 10^6	184.88 ± 26.10
p_5	-1.43 ± 1.50 x 10^7	-4.32 ± 0.26
p_6	-7.50 ± 4.16 x 10^7	-6.20 ± 2.20
p_7	-4.15 ± 0.091	-8.88 ± 1.59
p_8	-8.63 ± 2.00	-11.34 ± 7.54

Table 3. Estimated values of half-life ± standard deviation after parameter estimation.

Temperature [°C]	**5 optimal + 1 initial profiles**
48.0	1912.01 ± 574.2
50.0	1239.43 ± 266.8
52.0	800.68 ± 121.8
54.0	520.17 ± 57.42
56.0	342.81 ± 30.2
58.0	231.09 ± 17.4
60.0	160.53 ± 10.4

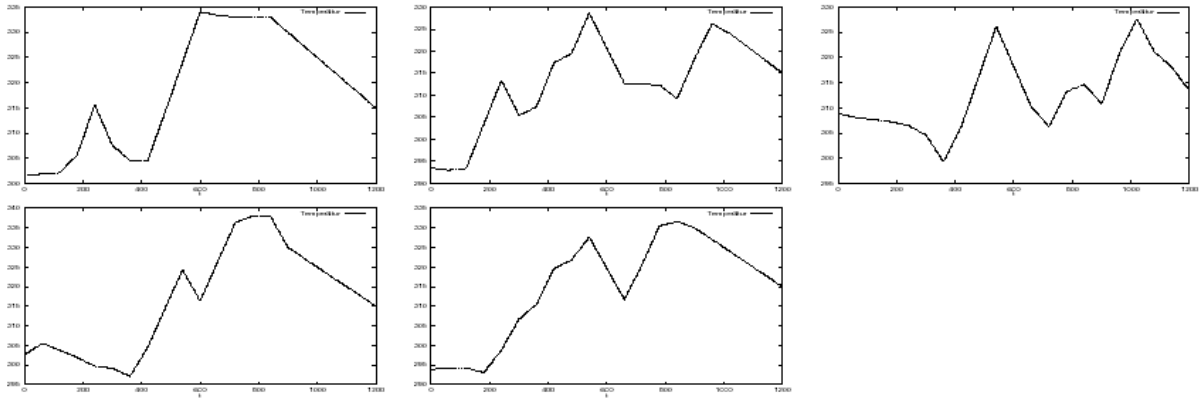


Figure 3. Optimal additional temperature profiles.

A)

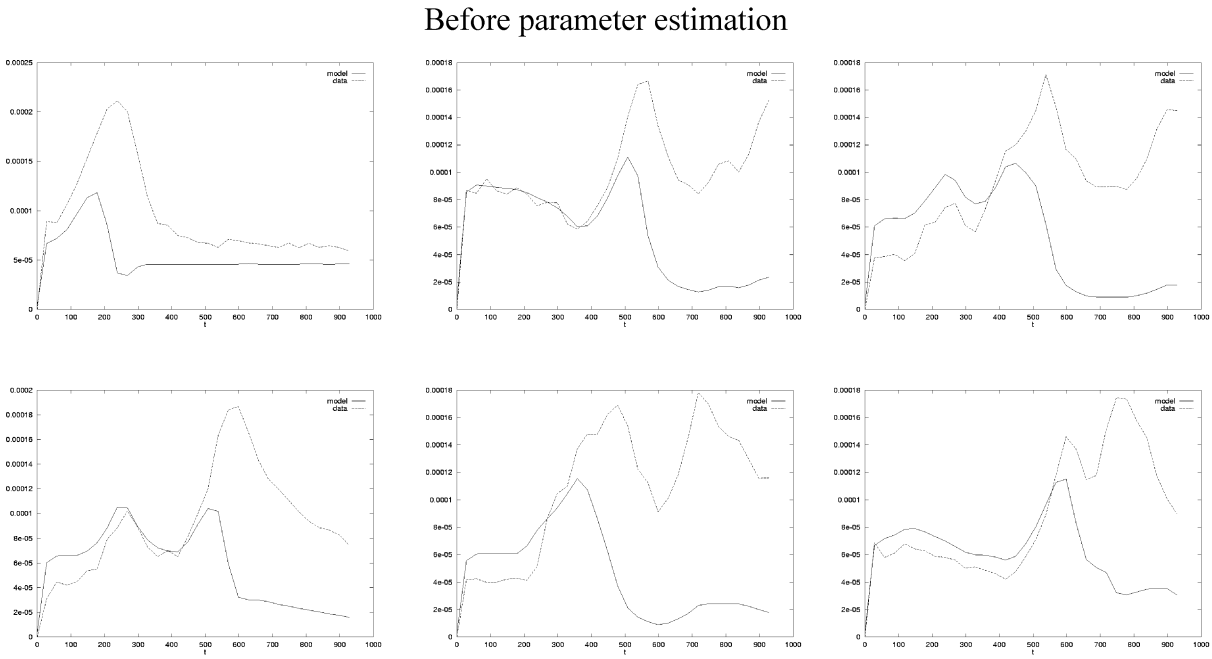


Figure 4A. Model (straigth line) vs. measurements (dashed line) before parameter estimation.

B)

After parameter estimation

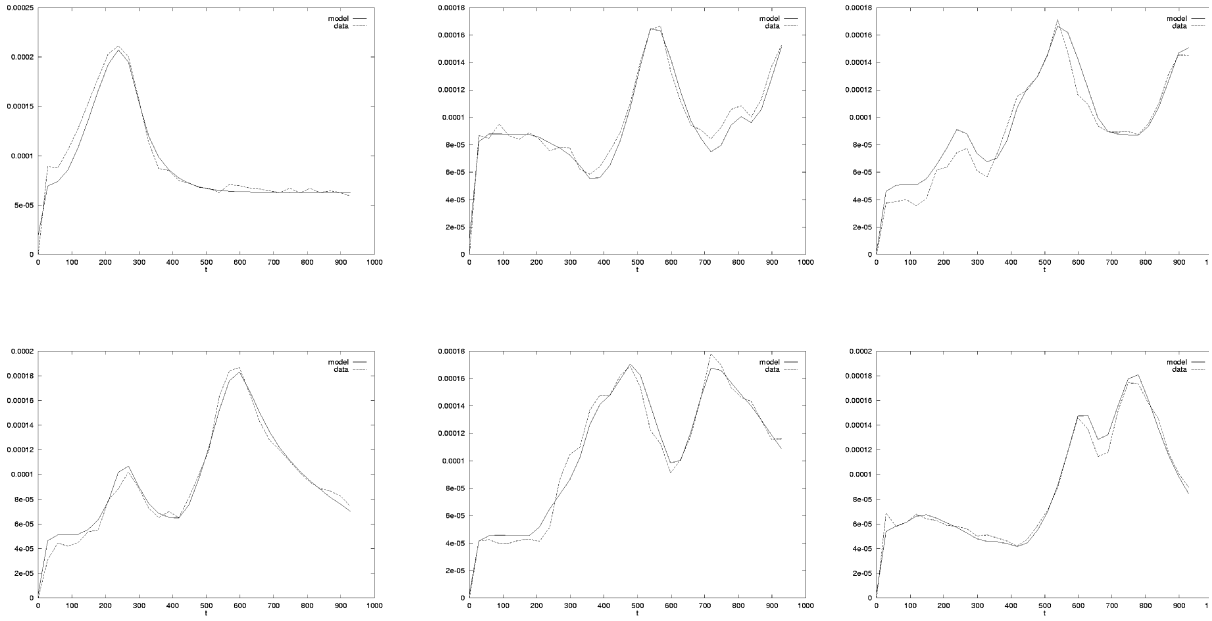


Figure 4B. Model (straigth line) vs. measurements (dashed line) after (**B**) parameter estimation.

DESIGN OF ROBUST OPTIMAL EXPERIMENTS OF NONLINEAR MODELS

As we can see the experimental design optimization problem (10), (11), (12) is formulated for the assumed parameter values which are, however, only known to lie in a possibly large confidence region. In this section we discuss how to construct robust experiments, that is experiments that are less sensitive to parameter uncertainty. We assume that the parameters are lying in the following ellipsoid around the nominal values of parameters p^0

$$E := E(\gamma, p^0) = \left\{ p : (p - p^0)^T \Sigma^{-1} (p - p^0) \leq \gamma^2 \right\}.$$

Here Σ is a positive definite matrix. To get a robust experimental design we formulate a worst-case problem, minimizing over the design variables ξ, the maximal value of $\phi(C)$ over the ellipsoid E

$$\min_{\xi \in \Omega} \max_{p \in E} \phi(C(\xi, p)). \qquad (13)$$

In this problem, we first consider constraints which do not depend on the parameters. The remaining (control) constraints are summarized by $\xi \in \Omega$. The optimization problem (13) is a semi-infinite programming problem. The solution methods for such problems require the determination of global optima of nonlinear subproblems [13] which may be computationally too expensive. In order to compute robust designs we suggest a modified approach. We apply Taylor expansion w.r.t. *p* to the min-max objective function:

$$\min_{\xi\in\Omega} \max_{p\in E} \left(\phi(C(\xi,p^0)) + \frac{\partial}{\partial p}\phi(C(\xi,p^0))(p-p^0)\right).$$

The inner problem is the maximization of a linear function subject to a convex quadratic constraint which can be solved explicitly:

$$\max_{p\in E} \left(\phi(C(\xi,p^0)) + \frac{\partial}{\partial p}\phi(C(\xi,p^0))(p-p^0)\right) = \phi(C(\xi,p^0)) + \gamma \left\| \frac{\partial}{\partial p}\phi(C(\xi,p^0))\Sigma^{1/2}\right\|_2.$$

This leads to the robust experimental design optimization problem

$$\min_{\xi\in\Omega} \left(\phi(C(\xi,p^0)) + \gamma \left\| \frac{\partial}{\partial p}\phi(C(\xi,p^0))\Sigma^{1/2}\right\|_2\right).$$

The second term in the cost function can be interpreted as a penalty for uncertainty in the parameters.

Parameter dependent constraints are treated in an analogous way. We substitute the nonrobust constraints

$$\psi(\xi,p^0) \le 0,$$

with the following constraints for the robust experimental design optimization problem.

$$\psi_R(\xi, p^0) := \psi(\xi, p^0) + \gamma \left\| \frac{\partial}{\partial p} \psi(\xi, p^0) \Sigma^{1/2} \right\|_2 \leq 0.$$

Numerical results

In this section we present numerical results on comparing robust and nonrobust designs for enzyme reaction kinetics. To illustrate the performance of the method, experiments for identification of 4 out of 8 parameters were designed in the sequential mode. The idea of the sequential approach is schematized in Fig. 5.

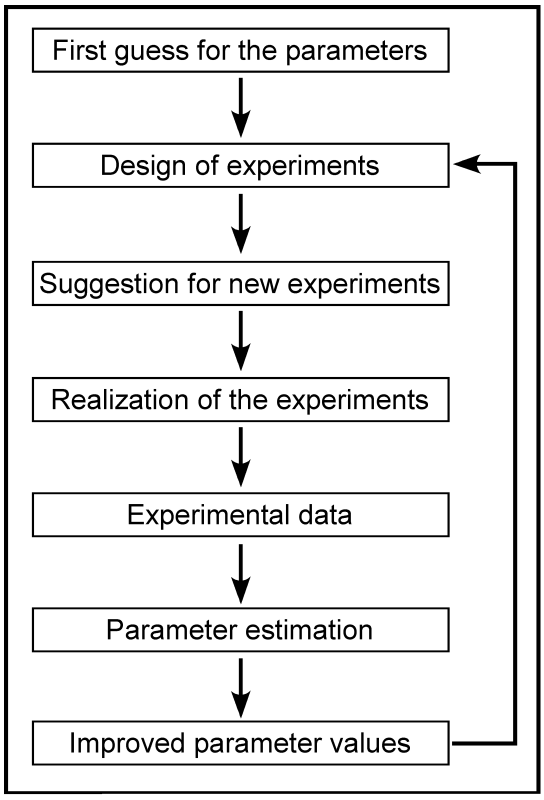


Figure 5. Sequential approach to experiment design.

For the "true" values of the parameter $p_1^{true} = 27.77, p_2^{true} = 50.15, p_3^{true} = 0.55, p_4^{true} = 185.25$ the normally distributed data were simulated using initial temperature profile as shown in Fig. 2.

The results of parameter estimation based on these data

$p_1 = 48.61 \pm 20.59$

$p_2 = 103.58 \pm 52.77$

$p_3 = 0.58 \pm 0.10$

$p_4 = 189.69 \pm 32.50$

were used for the computation of a nonrobust and a robust designs. The diagonal matrix with diagonal elements of the covariance matrix corresponding to the solution of parameter estimation problem was chosen as the matrix Σ in the robust design. Using the new optimized temperature profiles the data were simulated, parameter estimation was repeated with the new data and further new designs were constructed using the new estimates for the parameters. This procedure was repeated until we obtained trustworthy parameter estimates, see Table 4.

Table 4. Results of parameter estimation for initial design and 5 non-robust experiments (second column) and for initial desgin and 2 robust experiments (third column).

"true" values	$p_i \pm \sqrt{c_{ii}}$	$p_i \pm \sqrt{c_{ii}}$
$p_1 = 27.77$	27.44 ± 0.743	27.63 ± 1.266
$p_2 = 50.15$	49,33 ± 1.980	49.71 ± 3.330
$p_3 = 0.55$	0.620 ± 0.068	0.512 ± 0.072
$p_4 = 185.25$	209.85 ± 24.35	172.26 ± 24.41

The sequential design of robust experiments after 2 designs provide us with reliable parameter estimates, while we need 5 nonrobust experiments to get similar quality of the estimates. Although a computational time for computing a robust experiment is significantly greater than the time to compute a nonrobust experiment, clearly application of the robust sequential design makes it possible to reduce the number of real experiments and thus to reduce drastically experimental costs and the time necessary to identify the parameters.

CONCLUSIONS

Application of advanced methods for parameter estimation in dynamic processes governed by ODE and for design of robust optimal experiments allows reliable estimation of enzyme operating stability and the reduction of the number of experiments as well as experimental costs.

ACKNOWLEDGEMENTS

We thank the Deutsche Forschungsgemeinschaft (DFG) for financial support through SFB 359.

The part of the research was made in cooperation with Degussa AG.

REFERENCES

[1] Gupta, M.N. (1993) *Thermostability of Enzymes*. Springer, New York.

[2] Boy, M., Dominik, A., Vos, H. (1998) A method for fast determinationof biocatalyst process stability. *Chem. Engng Technol*. **21**: 570-575.

[3] Bommarius, A., Estler, M., Kluge, A., Werner, H., Vollmer, H., Bock, H.G., Schlöder, J.P., Kostina, E. (2001) Method to determine the process stability of enzymes. Patent Application EP 1 067 198 A1, European Patent Office, *Patentblatt 2*.

[4] Bock, H.G. (1981) Numerical treatment of inverse problems in chemical reaction kinetics. In: *Modelling of Chemical Reaction Systems*.(Ebert, K.H. Deuflhard, P., Jäger, W. Eds) Springer Series in Chemical Physics **18**, Heidelberg.

[5] Bock, H.G. (1987) Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen. *Bonner Mathematische Schriften* **183**: 264.

[6] Bock, H.G., Kostina, E., Schlöder, J.P. (2000) On the role of natural level functions to achieve global convergence for damped Newton methods. In: *System Modelling and Optimization: Methods, Theory and Applications.* (Powell, M.J.D., Scholtes, S. Eds) pp. 51-74. Kluwer, Boston.

[7] Stoer, J. (1971) On the numerical solution of constrained least squares problems. *Siam J. Numer. Anal.* **8**(2): 382-411.

[8] Griewank, A. (2000) *Evaluation Derivatives. Principles and Techniques of Algorithmic Differentiation*. Frontiers in Applied Mathematics. SIAM

[9] Bauer, I., Bock, H.G., Körkel, S., Schlöder , J.P.(1999) Numerical methods for initial value problems and derivative generation for DAE models with application to optimum experimental design of chemical processes. In: *Scientific Computing in Chemical Engineering II.* (Keil, F., Mackens, W., Voss, H., Werther , J.Eds) **2**, pp. 282-289, Springer, Berlin.

[10] Bauer, I., Bock, H.G. Körkel, S., Schlöder, J.P. (2000) Numerical methods for optimum experimental design in DAE systems. *J. Comput. Appl. Math.* **120**:1-25.

[11] Körkel, S.(2002) *Numerische Methoden für Optimale Versuchsplanungsprobleme bei nichtlinearen DAE-Modellen*. PhD thesis, Universität Heidelberg.

[12] Körkel, S., Bauer,I., Bock, H.G., Schlöder, J.P. (1999) A sequential approach for nonlinear optimum experimental design in DAE systems. In: *Scientific Computing in Chemical Engineering II.* (Keil, F., Mackens,W., Voss, H., Werther, J.,Eds) **2**, pp. 338-345, Springer, Berlin.

[13] R. Reemtsen, R., Rückmann, J.-J. (Eds.) (1998) *Semi-Infinite Programming. Nonconvex Optimization and its Applications*. Kluwer, Boston.